

Classroom Use of Test Accommodations: Issues of Access, Equity, and Conflation

By: [Jamie L. Schissel](#)

Schissel, J.L. (2014). Classroom use of test accommodations: Issues of access, equity, and conflation. *Current Issues in Language Planning*. DOI:10.1080/14664208.2014.915458

This is an Accepted Manuscript of an article published by Taylor & Francis Group in *Current Issues in Language Planning* on May 13 2014, available online at:

<http://www.tandfonline.com/10.1080/14664208.2014.915458>

*****© Taylor & Francis. Reprinted with permission. No further reproduction is authorized without written permission from Taylor & Francis. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. *****

Abstract:

Test accommodations are changes to test administration, responses, or the test itself that are offered to emergent bilingual students for standardized tests and also for classroom assessments in some states in the USA. Currently there is a lack of research examining the use of test accommodations as a pedagogical practice. This paper presents a study of the classroom use of accommodations using data from a comparative ethnographic case study of two different multilingual schools. A historical analysis and definitions of test accommodation effectiveness frame the analysis. Participant observation, field notes, interviews, and classroom artifacts are used to investigate how administrators, teachers, and students appropriate test accommodations as classroom practices and one class in particular serves as an entry point for unpacking classroom use of accommodations. The analysis positions test accommodations as a tool promoting increased access to assessments but cautions that this access may not be an equitable practice, especially when accommodation use conflates emergent bilingual students with students who have disabilities. The paper concludes by exploring the possible implications of using classroom practices to inform policies and the use of test accommodations for emergent bilinguals to explore more options for multilingual education and assessment.

Keywords: testing accommodations | English language learners | classroom assessment | bilingual education | language education policy

Article:

Introduction

Ms. Frederick's third grade class has finished their social studies unit on community services and now prepares to take the summative assessment. As students take out their pencils, Ms. Frederick tells the students that she can read test items aloud to students if

they are struggling with reading. As students complete the exam, they all raise their hands to have items read aloud; so many in fact that Ms. Frederick does not have time to go to a student each time a hand is raised. During the test, she not only reads items aloud, but also explains some of the questions and reminds individual students and the entire class about related previous class discussions. By the end of the test, she tells students to raise their hands for help rather than to have items read aloud. During the entire test, she spends most of her time with one student who has been in the United States for less than a year. As the class finishes up, the ESL teacher enters the room and takes this student aside to help her complete the test. Ms. Frederick and the rest of the students begin a new subject. (Field notes, 04.30.10)

In Ms Frederick's classroom and other classrooms in the USA that have English language learners receiving English as a second language (ESL) services, federal and state policies mandate that changes – referred to as *test accommodations* – are made to the test administration, test response, or the test itself for annual, standardized, high-stakes content tests. In this paper, I refer to English language learners or limited English proficient students – as they are termed in policies – as *emergent bilinguals*. The change in terminology foregrounds students' dynamic multilingualism or varying proficiency, over time, of home language(s) and English (García, 2009) and moves away from current biases that have removed the term bilingual from education policy in the USA (Crawford, 2002; Hornberger & Johnson, 2007; Johnson, 2010). Drawing attention to the dynamic multilingualism of emergent bilinguals is particularly relevant when discussing test accommodations because the use of accommodations aims to reduce construct-irrelevant variance due to English language proficiency, i.e. to allow emergent bilinguals to demonstrate their knowledge in a specific content area (e.g. mathematics and English language arts) rather than their unrelated English proficiency. In attempts to control for the influence of a developing proficiency in English impacting a test score, accommodations provide linguistically related supports such as the use of an interpreter or a bilingual word-for-word dictionary and linguistically unrelated supports such as small group administration with extra time.

Test accommodations are currently approved and allowed for use with emergent bilingual students receiving ESL services or students with disabilities. In the case of emergent bilinguals, states often mandate that test accommodations used for standardized testing must be implemented during classroom assessment and classroom instruction. Ms Frederick and her students live in a state that mandates the use of test accommodations in the classroom. This paper aims to unpack the interactions of testing policies and pedagogical practices by addressing the question *how are administrators, teachers, and students appropriating test accommodations as classroom practices?*

To explore test accommodation use, this paper first investigates historical approaches to standardized testing for emergent bilinguals. This section is followed by current definitions and research on test accommodation effectiveness. In the final section of the paper, I return to observations of Ms Frederick's classroom and three other classrooms drawn from a six-month

comparative ethnographic case study. This paper examines how the use of test accommodations for all students regardless of their eligibility can be framed as (i) promoting test accommodations as a tool for access rather than equity and (ii) conflating emergent bilingual students' needs with disabilities. The paper concludes by exploring the possible implications of using classroom practices to inform policies and the use of test accommodations for emergent bilinguals to explore more options for multilingual education and assessment.

Historical approaches to testing emergent bilinguals in the USA

Intelligence testing in the early 1900s represents the first instances of providing test accommodations to emergent bilinguals. These tests were administered using interpreters and alternative test forms. Test accommodations were used again in the 1970s with changes to achievement tests for students with disabilities in US schools, and in the 1990s for emergent bilinguals. This short overview will not go into the detail of some of the grossest misuses of these tests. Instead, it will focus specifically on the historical development of test accommodations practices with some discussion of the ideologies that have underpinned these practices.

Test accommodation practices were first used with the intelligence testing of immigrants at Ellis Island in 1913. The US government had contracted Henry Goddard to determine whether incoming immigrants had cognitive or developmental disabilities. Goddard and his colleagues administered a version of the Binet-Simon intelligence exam. Goddard had translated the exam from French to English and used it previously for children with developmental disabilities at the school where he was director (Spolsky, 1995). He and his team used this exam on Ellis Island with adult immigrants who had already passed through health and intelligence screening. When administering the test, Goddard hired trained interpreters to administer the test in Hungarian, Italian, and Russian and to Jewish immigrants – presumably interpreting into Yiddish, though the language was not specified. He used his results to claim that the highest mental age for these adults was 12 and that 79–83% of them were feeble-minded (the term used at that time) or developmentally disabled (Goddard, 1917). His findings were contested by his contemporaries, who also evaluated the intelligence of immigrants on Ellis Island, and more recently they have been criticized by modern-day researchers who have identified it as an extreme form of bias against a targeted group of ‘less desirable’ immigrants (see also Gelb, 1986). The use of interpreters as an accommodation remains in practice today in some states for emergent bilingual students.

In 1917, test developers took a different approach to the testing of emergent bilinguals. In the development of a large-scale intelligence test for the US military, Stanford University professor Lewis Terman altered the Binet-Simon intelligence scale to create the Stanford-Binet intelligence test. He used these scales to develop the Army Alpha and Beta exams that determined whether an incoming recruit had the potential to become an officer or should remain a soldier. Terman constructed a parallel test for emergent bilinguals and illiterate English

speakers. He claimed that this test, the Army Beta, was measuring intelligence in the same way that the Army Alpha did for literate English speakers. The Army Beta was administered with gestures and thus, the tasks did not require reading or oral language skills. Only a few utterances such as 'good' and 'hurry up' were to be spoken by test administrators in an attempt to eliminate language from the testing situation (Brigham, 1923).

Emergent bilinguals scored disproportionately low on the Army Beta in comparison with other groups of test takers on the Army Alpha and Beta; however, their scores increased in relation to the amount of time they had lived in the USA. When interpreting this trend in increasing scores, Brigham (1923) rejected the notion that increasing language or cultural knowledge with increasing length of residence may have been impacting on the scores. Instead, he used these data to posit that each new group of immigrants had increasingly lower intelligence. He cautioned that this would have implications for the country as a whole when immigrations were not only allowed into the USA, but also began to have children with other immigrants. He wrote that 'U.S. intelligence is declining, and will proceed with an accelerating rate as the racial admixture becomes more and more extensive' (p. 210).

Brigham's explanations were related to a larger movement of his time that found different ways to argue that language was not a factor in understanding intelligence because intelligence was an inherited trait (Hakuta, 1986). Brigham, Goddard, and Terman, among others, were strong supporters of this interpretation and – until the late 1920s – were actively involved in the Eugenics Movement, an explicitly racist and xenophobic movement that viewed many traits as inherited and often argued for selective breeding or sterilization to eradicate less desirable characteristics. Though it is not within the scope of this paper to discuss their position in detail, it is worth noting the ideological orientations of these researchers and how these connect to their interpretations of test scores.

Test accommodations akin to Goddard's use of interpreters and Terman's Army Beta were not formally instituted again until the 1970s. One reason for this gap is that, until the 1950s, challenges facing world economies precluded the development of many tests and immigration to the USA also dropped due to restrictive policies until the mid-1960s (Hing, 2004; Spolsky, 1995). During this time, the government also stopped testing immigrants on arrival in the USA and intelligence and achievement test use expanded in educational contexts. The 1960s and 1970s also saw a general trend toward trying to adopt more fair and equitable education and testing practices in US schools. There were court challenges, for example, addressing the misuse of intelligence tests scores from over to underrepresentation of emergent bilinguals in special education services (Baca & Cervantes, 1998).

Cummins addressed some of the possible reasons for the overrepresentation of emergent bilingual students as having disabilities in his research on how teachers and other test administrators interpreted language proficiency during intelligence and achievement testing. Cummins (1979, 1981) discussed the incongruence between teachers' observations about

students' proficiency in the dominant language (English) and students' performance on standardized and/or intelligence tests. In teachers' referrals of emergent bilingual students for special education services, they would often claim that a student could have a discussion or conversation and appeared fluent but then did not perform well on assessments. Cummins posited that the teachers were overgeneralizing students' perceived fluency from one-on-one encounters to all linguistic domains. This represents an example of the misinterpretation of language proficiency as a disability.

During this same time, test accommodations were recognized and supported at the national level as a practice for students with disabilities in different types of tests. Federal policies were enacted in 1973 and have continued to be renewed to ensure fair and equitable inclusion in testing for students with disabilities. When states exempted students with disabilities from testing rather than providing these test accommodations, the federal government stepped in to overrule these exemptions. The Americans with Disabilities Act in 1990 and the reauthorization of the Individuals with Disabilities Education Act (IDEA) in 1997 mandated that all states include students with disabilities in their assessment programs, with provisions for test accommodations being made available (Richards, 2003). In 2001, the *No Child Left Behind Act* (NCLB, 2002) expanded this mandate to include emergent bilinguals in assessments and to make the provisions for test accommodations available.

When NCLB was passed, most states did not have test accommodations for emergent bilinguals. They did, as previously mentioned, have test accommodations for students with disabilities. The majority of states applied these existing policies directly to emergent bilinguals with no differentiation. As recently as 2008, Rivera reported that 18 states did not distinguish their test accommodations policies for these two populations, essentially conflating the needs of emergent bilinguals with those of students with disabilities. In a close examination of these approved test accommodations, Abedi (2008) has explained that of the 73 test accommodations approved for emergent bilinguals, only 11–15% are appropriate for emergent bilinguals. In these states without differentiation of test accommodation practices, Abedi reported that some emergent bilinguals received the test accommodation of enlarged print, usually assigned to students with visual impairments.

New test accommodations proposed for use with the upcoming content tests from the two major testing consortia in the USA offer similar accommodations to those that have been available to emergent bilinguals since the early 1900s, such as alternative test forms in other languages and accommodations that were originally used for students with disabilities such as small group administration and extended testing time (Abedi & Ewers, 2013; PARCC, 2013).

This overview of the history of accommodation demonstrates that the linguistic needs of emergent bilinguals on these various forms of standardized tests both have and have not been included. Unpacking how research has begun to address test accommodation practices currently in use reveals another layer of the purposes for and uses of accommodations.

Definitions of test accommodation effectiveness

Researchers tend to define an accommodation as effective for a given population and test construct when that accommodation – or combination of accommodations – improves the score for the target population at a level of statistical significance ($p < .05$), while not affecting the score of the control group and also not altering the intended construct. For example, an accommodation such as reading items aloud is seen as effective if it raises the scores of emergent bilingual students compared to emergent bilingual students who did *not* have items read aloud and it does not change the score of non-emergent bilingual students who both *did and did not* have items read aloud. This accommodation additionally does not alter the construct being measured if it is offered for a math test, since the test is not measuring reading comprehension. However, the same accommodation is deemed ineffective for a language arts test because that test is measuring reading comprehension.

When evaluating the effectiveness of test accommodations, the connection of the intended purpose of reducing construct-irrelevant variance due to language with the construct validity and consequential validity of the test is worth emphasizing. Construct validity relates to the breadth and depth of the test measuring what it is intended to measure, whereas consequential validity is connected to stakes or consequences attached to test use. Messick (1989) argued that in addition to the traditional evidential approaches to ascertaining construct validity, test developers also needed to understand the consequential validity of using tests in relation to construct underrepresentation and construct-irrelevant demands of an assessment. When testing emergent bilinguals in content areas, addressing the possible construct-irrelevant variance due to the English demands of the test is essential in order to ensure that the inferences from the scores are meaningful. In the current high-stakes testing situation in the USA, the consequences attached to poor performance on external assessments (e.g. restricted funding, state control of schools) make the issue of test accommodation effectiveness especially important.

Current research on the effectiveness of test accommodations for emergent bilinguals has largely been focused on large-scale standardized assessments such as the National Assessment of Education Progress. Effectiveness of accommodations has been most consistently found in research on the use of English glossaries. This accommodation has shown statistical significance in improving test scores and is most beneficial when combined with extra time on math and science large-scale assessments (Abedi, 1999; Abedi, Courtney, Mirocha, Leon, & Goldberg, 2005). Extra time, the most frequently used accommodation (Rivera & Collum, 2006), has increased scores for emergent bilingual students at a level of significance on math and reading comprehension assessments (Hafner, 2001).

For accommodations that incorporate other languages, the evidence has been limited. In terms of multilingual support, providing first language translations in the margins in science tests has shown evidence of effectiveness (Abedi, 2001). In addition, Levin, Shohamy and Spolsky (2003, as cited in Shohamy, 2006, 2011) have shown how bilingual testing in Hebrew and Russian

rather than monolingual testing improved the performance of Russian immigrants in Israel. Home language assessments for students whose medium of instruction was the home language also yielded significantly better performance (Hofstetter, 2003; Menken, 2008).

The linguistic complexity of test items has been shown to impact on emergent bilingual students' performance (Martinello, 2008). Linguistic simplification, or a reduction in the lexical, syntactic, discourse, and presentation complexity has been characterized as not only a valid accommodation, but also one that is feasible in terms of a school's resources (Rivera & Stansfield, 2001; Stansfield, 2002). There is a small amount of evidence that shows that linguistic simplification is beneficial for emergent bilingual students (Abedi et al., 2005) while having no effect on the control group (Rivera & Stansfield, 2001). The availability of this accommodation, however, is currently limited.

Reports have cautioned that the research on the effectiveness of accommodations remains limited (see, e.g. Kieffer, Lesaux, Rivera, & Francis, 2009; Koenig & Bachman, 2004; Pennock-Roman & Rivera, 2007) and thus the inferences made from the scores of content assessments taken by emergent bilinguals also remain, to some degree, circumspect. Despite these limitations, the new standardized summative assessments being developed by the two testing consortia, PARCC (Partnership for Readiness for Colleges and Careers) and Smarter Balanced Assessment Consortia and set for implementation for the 2014–2015 academic year will continue to use test accommodations as the most dominant method to address the needs of emergent bilinguals. These accommodations include word-to-word English and bilingual dictionaries for non-construct-related vocabulary, extended time, and small group administration among others (Abedi & Ewers, 2013; PARCC, 2013).

The issues surrounding test accommodation use is exacerbated by state mandates that require the use of test accommodations not only during standardized testing where there is limited evidence of effectiveness, but also during classroom assessment and instruction where there is little to no evidence of their use (Schissel, 2013). The conclusion by Kieffer et al. (2009) that 'accommodations are not a solution to the larger issues of promoting the academic skills of ELLs [emergent bilingual students]' (p. 1190) frames this investigation of how administrators, teachers, and students appropriate the policies of test accommodation during classroom assessment as a possible way to see how classroom use of test accommodations and other multilingual scaffolds may expand practices during standardized testing.

Test accommodations in the classroom

To understand test accommodation use, this study examines data from administrators, teachers, and students' appropriation of test accommodations that were collected from January to June in 2010 in two urban schools in Pennsylvania. The comparative ethnographic case study examined one 3rd and one 5th Grade classroom in two multilingual schools. All names of the participants, including the school names, are pseudonyms. Olga Nolla Charter School was a kindergarten to

8th Grade school that was introducing a Spanish and English bilingual program. At the time of the study, kindergarten through 2nd Grade classes were participating and the rising 3rd Grade class was set to begin participation in the following academic year. The students in the school in all grades had varying degrees of exposure to and proficiency in Spanish, as did the faculty and staff. Westerville Elementary School was a 1st to 5th Grade school with English-only instruction. The students came from a variety of immigrant backgrounds and represented 12 different languages in the two classrooms of the study, 30 in the school, and over 70 in the district. The teachers and staff were primarily English speakers.

The opening vignette from Ms Frederick's 3rd Grade classroom occurred at Westerville Elementary School and serves as the foundation for this analysis of test accommodations as pedagogical practices. For data collection, I visited the schools on alternating weeks and spent the entire school day with the class for two to three days per week. Field notes based on participant observation served as the primary means of data collection. Recorded semi-structured interviews were conducted with the teachers and administrators at the schools as well as follow-up interviews with one district-level official were conducted for each school district approximately two years after the initial data collection. Informal, non-recorded interviews also occurred with the teachers and students. Classroom artifacts, such as pictures of the rooms throughout test preparation and of student work were taken in addition to the collection of publically available documents and data on the schools policies, test performances, and other information about the general background of the schools and districts. Data events were classified as occurring during practice standardized testing, classroom assessment, and classroom instruction situations. Themes emerged through coding of how test accommodations were initiated, responded to, and evaluated and analytic memos served as initial entry points for data analysis.

In this paper, the historical development of test accommodations and definitions of test accommodation effectiveness serve as lenses for two interpretations of how test accommodations were appropriated in classrooms. These analyses pointed to the use of test accommodations for all students as providing access to testing and as conflating emergent bilinguals with students with disabilities. Before unpacking these analyses, Ms Frederick's administration of the social studies test is presented in more detail.

Test accommodation in the classroom: returning to Ms Fredrick's class

The vignette presented at the beginning of this paper highlights some of the used policy-approved test accommodations for standardized testing situations such as reading test items aloud in addition to practices that are not policy-approved test accommodations, such as referring to previous class discussions. Ms Frederick's administration of this social studies test also serves as an entry point for examining how administrators, teachers, and students appropriated these policies in ways that aligned with and diverged from current mandates.

The policies for test accommodation availability for emergent bilinguals stipulate that only students who are receiving ESL services are eligible for accommodations. However, Ms Frederick as well as other participants in this study emphasized that they found accommodations to be beneficial for all students. In Ms Frederick's classroom, 7 of her 19 students were eligible for accommodations. Students who were recently exited from ESL services were also eligible at the discretion of teachers and administrators. The five students in this class who had been recently exited from ESL services had been deemed able to take the assessments without accommodations. Students with disabilities can also receive this same test accommodation of reading test items aloud. In this classroom, one student was classified as both receiving special education and ESL services. Another student was being evaluated for special education services but the evaluation was not complete at the time of this test administration.

The mostly teacher-created assessment covered explicit content that was included in the class. At the beginning of the test administration, Ms Frederick initiated the use of the accommodation to read test items aloud to students. This accommodation is permitted for the math and science standardized annual exams. During the test administration, she not only offered accommodations to all students, but all students responded and requested the accommodation multiple times. Ms Frederick was not able to provide accommodations for every request. During the teacher and student interactions throughout the administration, Ms Frederick deviated from reading questions aloud to offering non-policy approved accommodations. She contextualized the test content by reminding students of stories that they had read in class and past class discussions. For two specific items that students struggled with, she offered two different strategies. For an open-ended question that asked about community services, she rephrased and simplified the prompt. For a multiple-choice question, she referenced a previous conversation in class.

While students were finishing, Ms Frederick reminded the students to check their answers, which was a common strategy used by teachers in this study for all students when they are taking standardized tests. She also offered a revised or modified accommodation. At the onset of the test administration, Ms Frederick stated that she could read items aloud to students. However, she offered other accommodations beyond reading the questions aloud. When she stated in the final minutes of the assessment 'If you need help raise your hand,' she vocalized this change in her approach to accommodations.

When she collected the tests and transitioned to the next activity, there was little evaluation of the students' work. Evaluation of students' responses to accommodations came mostly in non-verbal communication, such as lingering near a student as they answered a question, and leaving after they had chosen the correct response. During these interactions, the teacher and student were able to enter into a dialog about the test item. During this time, the ESL teacher, Ms Ruby, provided the policy-approved accommodation of small group test administration.

Test accommodations as a tool of access

Test accommodations were positioned in this and other interactions as a beneficial practice that promoted equity and access of assessments for emergent bilinguals. During a formal interview, Ms Frederick explained how she met the unique needs of each student in her class; for example, she made sure to seat students in areas where they could stay focused, she gave students individual attention, and she checked that one student had eaten breakfast. Her care and attention to her students is evidence of an earnest investment in the students and their academic trajectories (04.30.10).

With assessments, her recognition of students' individual needs was trumped by her overarching view that accommodations were a beneficial tool that therefore should be given to all students. As she explained, 'If I'm going to make accommodations for some, I'm going to make them for all because it's not going to be hurtful.' When asked directly if she thought they were helpful, she responded quickly 'yeah, yeah'. The use of accommodations had also become habitual; according to Ms Frederick 'I kind of just do it all the time because I'm so used to it' (interview, 04.30.10). This perspective of providing the same testing accommodations for all students was brought to bear in her practice.

Looking at Ms Frederick's actions during the classroom assessment situation and her responses to the interview questions offers rich insights into her use of accommodations specifically during classroom assessment situations. She offered the same accommodations to all students but the implementation of accommodations for all students differed based on her availability. Although she offered accommodations to all students, she spent the majority of her time with one student who had recently immigrated to the USA. In the classrooms in Olga Nolla Charter School, accommodations such as interpretation from English to Spanish were also offered to all students but rather than all students requesting the accommodations, reports from a survey of test accommodation use during standardized testing and participant observation field notes document that few students requested this accommodation. Those who requested interpretation did so at the beginning of testing or classroom instruction and that these few requests diminished to zero after the first 5–10 minutes.

Teachers and administrators expressed the belief that accommodations were helpful across research sites. Mr Schuhmacher, a recently retired administrator from Olga Nolla's school district accountability office, expressed his opinion about the benefits of accommodations thus: 'I don't think the accommodations we've talked about and the ones approved by [the Department of Education] give students an unfair advantage. I think they do level the playing field, which is their intention' (interview, 02.15.12). He agreed with Ms Frederick that accommodations were helpful and went one step further to explain that, although they were helpful, they offered opportunities for students to show what they know and did not artificially raise scores. He also used the phrase 'level the playing field' to explain how accommodations function. The idea that accommodations can make all things equal as implied with this phrase and that accommodations are helpful reinforced the idea that test accommodations created equitable practices.

Vice Principal Morena and English as a second-language program coordinator, Mr Nuñez at Olga Nolla Charter School explained how they found accommodations to be beneficial for all students. Mr Nuñez began by answering the question about the effectiveness of accommodations and Ms Morena followed his response by explaining how they now have applied these accommodations to all students.

Yes [accommodations are helpful] because it's a reduced set-up. Sometimes you have eight students, sometimes you have only three students so they are completely out of the whole environment with other students where there are more distractions. [...] So reduced groups help a lot. (Nuñez interview, 05.19.10)

[...] We make accommodations actually for all our students. [...] So we've learned that with children not having a full class testing at one time or one area is not the best idea because they get distracted. So we've created for our regular, whatever that might be, regular students we've created that there's no more than seven to ten students in a classroom testing in one classroom tops. So we've creating an environment where it's a smaller testing environment so there're less distractions. (Morena interview, 05.19.10)

In answering the question about the effectiveness of accommodations, Mr Nuñez explained that he felt they were helpful and that this was especially true when tests were administered in small groups. Ms Morena expanded on Mr Nuñez's response by explaining that because of the benefits that they have seen with other students, they now use small groups for all students. Ms Morena supervised kindergarten through 3rd Grade so it was possible that by all students, she was talking about this group within Olga Nolla. One aspect of determining test accommodation effectiveness focuses on this type of expanded use, that an accommodation designed for students receiving ESL services should not impact on the scores of students not receiving ESL services who use this same accommodation. Thus, this understanding of accommodation use for all students would be defined as a fair and equitable practice for both emergent bilinguals and non-emergent bilinguals because it would not negatively impact either groups' scores. But viewing accommodations as equitable through this definition only stands up if the test accommodations used had indeed been shown to be effective for emergent bilinguals.

Closer examination, however, points to test accommodations being more of a practice of access than equity for emergent bilinguals. Similar to test accommodation use with the Army Beta, accommodations both then and now have been used to increase the inclusion of emergent bilinguals in testing. In their appropriation of these policies, teachers and administrators viewed accommodations as beneficial and a method to support the inclusion of emergent bilinguals in accountability systems on a level playing field, so to speak.

Issues of fairness or equity also relate to the effectiveness of test accommodations. In these different instances of inclusion with the Army Beta and current test accommodations, the inferences made from the test scores were seen to represent the test takers' achievement. The

flawed interpretations of the Army Beta have been recognized (Hakuta, 1986) and, despite the critiques of the scores by emergent bilinguals today using test accommodations (Kieffer et al., 2009; Koenig & Bachman, 2004), policies continue to view these scores as providing valid inferences on emergent bilingual students' learning. The accommodation that Ms Frederick was using in the classroom, i.e. reading test items aloud, does not have consistent evidence supporting it as an effective accommodation for emergent bilinguals. Ms Frederick's use of this accommodation alone would most likely not have benefited her emergent bilingual students.

Instead, her use of pedagogical practices during testing sheds light on some of the discrepancies between test administration of standardized tests and classroom assessments and the view of test accommodations as beneficial. Her scaffolding approach that drew from a range of pedagogical practices possibly allowed her to make the test more equitable for her emergent bilingual students. Ms Frederick, the other teachers and administrators' expanded use of test accommodations beyond state mandates may also contribute to their positive view of test accommodations. Had she and the other participants strictly implemented test accommodations during classroom assessments as mandated in the state-policy, perhaps her view of test accommodations would have differed.

Though test accommodations seemed to be positioned as a tool of equity and access for emergent bilinguals, these data point only to accommodations as a tool of access, with the question of equity left unanswered. To understand issues of equity, the following section examines the decisions behind test accommodation use and how the needs of emergent bilinguals were equated with the needs of students with disabilities.

Test accommodations conflating the needs of emergent bilinguals

Despite the limited evidence on the effectiveness of accommodations, the purposes behind their use are clear in addressing an emergent bilingual student's linguistic needs. As previously stated, test accommodations are designed to reduce construct-irrelevant variance due to language. Test accommodations used with students with disabilities are similarly designed to address the specific needs of those students. The use of accommodations with students who are not receiving ESL services or do not have disabilities would have little to no impact because, as discussed, it does not alter the construct being measured. Using an accommodation designed for students with disabilities with emergent bilinguals or vice versa would similarly not impact on the performance of either group because for an accommodation to be effective, it needs to be appropriately assigned to a student based on his or her individual needs. This introduces a caveat in the above approach of using the same accommodations for all students if multiple students' needs have to be addressed with multiple accommodations. Unpacking explanations as to why an accommodation was used with the students in this and other classrooms can highlight some possible misunderstandings of the purposes of test accommodations that may lead to misuse.

Examining how test accommodations are chosen demonstrates how test accommodation use can begin to blur the lines between student categories. Instead of viewing accommodations as a tool to meet the specific needs of students, Ms Frederick and other teachers and administrators focus on test accommodations as a tool to help improve students' performance in exams. In explanations for using test accommodations for all students, emergent bilinguals included, they conflate the needs of emergent bilinguals with other student groups, most specifically students with disabilities.

In using test accommodations with all students, students that Ms Morena defined as, 'regular, whatever that may be, regular students', she questioned the different categories into which students in Olga Nolla were placed. Rather than viewing emergent bilinguals who are receiving ESL services as different from the emergent bilinguals in the classroom who are not receiving ESL services, she seems to challenge the ambiguity of the official boundaries. Other participants in this study also blurred these lines of student categorization in discussions about test accommodations. At Olga Nolla, the emergent bilingual students receiving ESL services and who also had disabilities were most often only given accommodations to address their disability rather than their language needs. Most often, teachers and administrators seemed to conflate the needs of emergent bilinguals with those of students with disabilities, either explicitly by stating that emergent bilinguals were treated the same as students with disabilities or more subtly by treating the two groups of students as having the same needs for test accommodations.

Ms Frederick was the most explicit in this conflation. Her questioning of the categories of students was based on her view that students' achievement was low overall. 'I don't really change anything for the ELLs [emergent bilinguals receiving ESL services] other than the same accommodations I would do for somebody with an IEP [individualized education program, for a student with disabilities], same type deal.' When asked to explain her use of accommodations further, she explained, 'They're [the students in Ms Frederick's class] low. The whole group level is low. Like even the smartest kids would only be average in a normal classroom setting' (interview, 04.30.10). Specifically for emergent bilingual students receiving ESL services, Ms Frederick reported that she provides them the same accommodations as she provides for students with disabilities. In her comments she demonstrates how, because she viewed her entire class as low achieving, she used strategies that she knew that worked for students with disabilities and for all students. Her choices and characterization of her class is noteworthy, as all but one student in this class spoke a language other than English at home, which would make the linguistic needs prominent for all students.

School and district administrators also addressed this possible conflation of emergent bilingual students' needs with students with disabilities. Former district administrator, Mr Schuhmacher also seemed to group emergent bilinguals with students with disabilities when discussing the purposes for using test accommodations.

Everyone wants students to do as well as they can do and clearly that gives it another oomph to make sure that if a student is entitled to an accommodation whether an ELL [emergent bilinguals receiving ESL services] or students with disabilities, that they get the accommodation. (Interview, 02.15.12)

Throughout his interview, he did not differentiate accommodations for emergent bilinguals and students with disabilities, thus obscuring the different approaches of accommodations for these different students' needs. Ms Morena offered some clarification that different accommodations were made available for these different students' needs.

We make accommodations actually for all our students. There are the accommodations that we make in-house that we are mandated to do for ESL students. There are also others we have to make for special education students but we like to provide our students with the best environment possible. (Interview, 05.19.10)

These data do not necessarily mean that teachers and administrators felt that these groups have the exact same needs. As shown, Ms Morena offered some distinction. Mr Schuhmacher's lack of differentiation can be interpreted as a less explicit form of blurring the boundaries between these student groups. Ms Frederick's statements offer the clearest evidence that there exist views that explicitly connect these groups by offering the 'same type deal' for emergent bilingual students as she would for a student with a disability.

These discourses and actions conflating emergent bilingual students and students with disabilities may be difficult to avoid due to the influence of how test accommodations have been articulated in education policy. As mentioned previously, test accommodation policies originated in 1973 specifically for students with disabilities. Present constraints in Title I of NCLB (2002) state that accommodations must be made available for students with disabilities and emergent bilingual students in accordance with the IDEA, a policy that specifically addresses the needs of students with disabilities. These policy issues may also contribute to the ways in which the needs of these two groups are not fully differentiated. This conflation could also be connected to previous deficit views that positioned bilingualism as a handicap in the 1950s (Hakuta, 1986).

In relation to test accommodation effectiveness, this conflation inhibits the use and development of possible test accommodations that could be beneficial for emergent bilingual students. Dismissing this phenomenon as a reflection of policies and practices for these two groups to receive accommodations does not meet the needs of either population. Creating policies and practices of test accommodations under a new name such as multilingual scaffolding for emergent bilinguals may be one way to emphasize the distinct needs of these student groups.

Discussion and implications

The study presented here has highlighted some of the ways in which administrators, teachers, and students appropriate test accommodation policies. The use of accommodations in Ms

Frederick's classroom highlighted the complexities of accommodations at the classroom level and how test accommodation policies often do not align with pedagogical practices. The history and overview of research on effectiveness point to some of the limitations of what is known about test accommodations. Using these lenses to look closely at Ms Frederick's classroom assessment and the views of other participants in this study, the analysis concludes that test accommodations can be a tool for promoting increased access to assessments but cautions that this access may not be an equitable practice, especially when accommodation use conflates the needs of emergent bilingual students with those of students with disabilities. These findings, however, may offer a window into seeing how classroom practices can inform the use of test accommodations for emergent bilinguals to foster more equitable multilingual education and assessment practices.

The evidence presented in this paper has focused on an expanded use of test accommodations for all students and changes in accommodations from policy-approved accommodations such as reading test items aloud to helping students through rephrasing prompts and contextualizing test content. What could not be ascertained from these data are the potential implications for student learning if a teacher were to implement test accommodations with fidelity during classroom assessment and instruction or the impact of using these classroom practices during standardized testing.

The example from Ms Frederick's class could be used to highlight how the current pedagogical practice can be integrated into assessment procedures. Further, her use of test accommodations for all students and the support of that practice by other participants in this study make the case for removing restrictions on the use of accommodations for only certain groups of students. If, for example, small group administration does not impact on the test construct, all students could be given this accommodation. This use of an accommodation for all students could support the addition of flexible testing practices and policies that recognize the variety of student needs in a classroom.

The approaches that teachers already take in the classroom to adapt content and instruction to emergent bilingual students' specific educational and linguistic needs can serve as a starting point to diversify the current available test accommodation practices. Different forms of multilingual scaffolding for emergent bilinguals could also be used. Exploring how emergent bilinguals use different multilingual and multimodal resources in the classroom and how these may be integrated into the test situation is another possibility (see Logan-Terry & Wright, 2010; Shohamy, 2011). Taking this grounded approach can expand test accommodations into a new territory where teaching is used to influence testing.

References

1. Abedi, J. (1999). *NAEP math test accommodations for student with limited English proficiency*. Paper presented at the annual meeting of the American Educational Research Association, Quebec, Canada.
2. Abedi, J. (2001). *Validity of accommodations for English language learners*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
3. Abedi, J. (2008). Utilizing accommodations in assessment. In E. Shohamy & N. Hornberger (Eds.), *The encyclopedia of language and education volume 7: Language testing and assessment* (pp. 331–348). New York, NY: Springer Science and Business Media LLC.
4. Abedi, J., Courtney, M., Mirocha, J., Leon, S., & Goldberg, J. (2005). *Language accommodations for English language learners in large-scale assessments: Bilingual dictionaries and linguistic modification* (CSE Report 666). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
5. Abedi, J., & Ewers, N. (February, 2013). Accommodations for English language learners and students with disabilities: A research-based decision algorithm. *Smarter Balanced Assessment Consortium*. Retrieved from <http://www.smarterbalanced.org/wordpress/wp-content/uploads/2012/08/Accomodations-for-under-represented-students.pdf>
6. Baca, L. M., & Cervantes, H. T. (1998). *The bilingual special education interface*. Upper Saddle River, NJ: Prentice Hall.
7. Brigham, C. C. (1923). *A study of American intelligence*. Princeton, NJ: Princeton University Press.
8. Crawford, J. (2002). *Obituary: The Bilingual Education Act, 1968–2002*. Retrieved from http://www.languagepolicy.net/books/AEL/Crawford_BEA_Obituary.pdf
9. Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, 19, 121–129.
10. Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). Los Angeles: Evaluation, Dissemination and Assessment Center, California State University.
11. García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Oxford: Wiley-Blackwell.
12. Gelb, S. (1986). Henry H. Goddard and the immigrants, 1910–1917: The studies and their social context. *Journal of the History of the Behavioral Sciences*, 22, 324–332.
doi: 10.1002/1520-6696(198610)22:4<324::AID-JHBS2300220404>3.0.CO;2-Y

- 13.** Goddard, H. H. (1917). Mental tests and the immigrant. *Journal of Delinquency*, 2, 243–277.
- 14.** Hafner, A. L. (2001). *Evaluating the impact of test accommodations on the test scores of LEP students and non-LEP students*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- 15.** Hakuta, K. (1986). *The mirror of language: The debate on bilingualism*. New York: Basic Books.
- 16.** Hing, O. B. (2004). *Defining America through immigration policy*. Philadelphia, PA: Temple University Press.
- 17.** Hofstetter, C. (2003). Contextual and mathematics accommodation test effects for English-language learners. *Applied Measurement in Education*, 16(2), 159–188.
doi: 10.1207/S15324818AME1602_4
- 18.** Hornberger, N. H., & Johnson, D. C. (2007). Slicing the onion ethnographically: Layers and spaces in multilingual language education policy and practice. *TESOL Quarterly*, 41(3), 509–532. Special Issue on Language Policy.
- 19.** Johnson, D. C. (2010). Implementational and ideological spaces in bilingual education language policy. *International Journal of Bilingual Education and Bilingualism*, 13(1), 61–79.
doi: 10.1080/13670050902780706
- 20.** Kieffer, M. J., Lesaux, N. K., Rivera, M., & Francis, D. J. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research*, 79(3), 1168–1201. doi: 10.3102/0034654309332490
- 21.** Koenig, J. A., & Bachman, L. F. (Eds.). (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessment*. Washington, DC: National Academies Press.
- 22.** Logan-Terry, A., & Wright, L. (2010). Making thinking visible: An analysis of English language learners' interactions with access-based science assessment items. *AccELLerate!*, 2(4), 11–14.
- 23.** Martinello, M. (2008). Language and the performance of English language learners in math word problems. *Harvard Educational Review*, 78, 333–68.
- 24.** Menken, K. (2008). *English learners left behind: Standardized testing as language policy*. Clevedon: Multilingual Matters.
- 25.** Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* 3rd ed (pp. 13–103). New York: Macmillan.

26. No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107–110, § 115, Stat. 1425 (2002).
27. PARCC. (April, 2013). *Draft: PARCC test accommodations manual*. Retrieved from <http://ca539dfd55636c55e922fd4c048d1c793e15a27f954b34a49d25.r49.cf1.rackcdn.com/PARCCDraftAccommodationsManualforSWDEL.pdf>
28. Pennock-Roman, M., & Rivera, C. (2007). *The differential effects of time on accommodated vs. unaccommodated content assessments for English language learners*. Center for Assessment Reidy Interactive Lecture Series.
29. Richards, D. M. (2003). An overview of §504. *Council of Educators for Students with Disabilities*. Retrieved from http://www.504idea.org/504_Overview_Fall_2003.pdf
30. Rivera, C. (2008, January 15–16). *Defining and refining accommodations appropriate for English language learners*. Paper presented at the Addressing achievement gaps: Language acquisition and educational achievement of ELLs conference hosted by the Educational Testing Service.
31. Rivera, C., & Collum, E. (2006). *State assessment policy and practice for English language learners: A national perspective*. Mahwah, NJ: Lawrence Erlbaum.
32. Rivera, C., & Stansfield, C. W. (2001). *The effects of linguistic simplification of science test items on performance of limited English proficient and monolingual English-speaking students*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
33. Schissel, J. L. (2013). *Policy-enforced washback: Legislating the influence of testing on teaching*. Work-in-progress paper presented at the annual meeting of the American Association of Applied Linguistics, Dallas, TX.
34. Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. New York, NY: Routledge.
35. Shohamy, E. (2011). Assessing multilingual competencies: Adopting construct valid assessment policies. *The Modern Language Journal*, 95(3), 417–429. doi: 10.1111/j.1540-4781.2011.01210.x
36. Spolsky, B. (1995). *Measured words: The development of objective language testing*. Oxford: Oxford University Press.
37. Stansfield, C. W. (2002). Linguistic simplification: A promising test accommodation for LEP students? *Practical Assessment, Research & Evaluation*, 8(7). Retrieved from <http://PAREonline.net/getvn.asp?v=8&n=7>